

СИСТЕМА ПОСТРОЕНИЯ РУССКОЯЗЫЧНЫХ ЛЕКСИЧЕСКИХ РЕСУРСОВ НА ОСНОВЕ ЛЕКСИЧЕСКИХ ЗАМЕЩЕНИЙ

Усталов Д.А.^{1*}, Игушкин С.А.²

¹⁾ Уральский федеральный университет имени первого Президента России Б.Н. Ельцина, г. Екатеринбург, Россия

²⁾ Университет ИТМО, г. Санкт-Петербург, Россия

*E-mail: dmitry.ustalov@urfu.ru

A SYSTEM FOR CONSTRUCTING RUSSIAN LEXICAL RESOURCES USING LEXICAL SUBSTITUTIONS

Ustalov D.A.^{1*}, Igushkin S.A.²

¹⁾ Ural Federal University, Yekaterinburg, Russia

²⁾ ITMO University, Saint Petersburg, Russia

Annotation. In this study, we have adopted the lexical substitutions workflow for producing thesauri by transforming open questions into multiple choices. The approach has been implemented as a mobile application. The reported study was funded by RFBR according to the research project No. 16-37-00354 мол_а.

Краудсорсинг является эффективным способом построения языковых ресурсов, нашедшим своё применение при разработке баз знаний, снятии семантической неоднозначности, а также при решении других задач создания структурированных данных и формализации человеческих знаний [1]. Одна из тенденций современных исследований краудсорсинга состоит в повышении доступности и скорости разметки путём создания простых в использовании мобильных интерфейсов для современных смартфонов и планшетов [2].

Проект Yet Another RussNet (<https://russianword.net/>), разрабатываемый с 2013 г., направлен на построение современного электронного тезауруса русского языка при помощи краудсорсинга [3]. В процессе выполнения проекта создан инструмент построения и редактирования синонимических рядов, предназначенный для работы специалистов-лексикографов и студентов-лингвистов. Важным направлением исследований является расширение целевой аудитории и привлечение волонтеров к работе над материалами тезауруса, для чего была адаптирована методика краудсорсинга на основе явления лексических замещений, состоящая из этапов поиска замещений, выравнивания смыслов и сопоставления значений [4]. Адаптация методики производится путём замены заданий с открытым вопросом на задания со множественным выбором для снижения количества потенциальных ошибок и повышения доступности заданий для волонтеров, не являющихся специалистами в области языкознания.

В настоящее время разработана мобильная система построения тезауруса на основе лексических замещений. Система реализована с использованием трёх-звенной архитектуры «клиент–логика–данные» (рис. 1). Клиентский слой обес-

печивается мобильным приложением конечного пользователя. Логический слой обеспечивается удалённым сервисом управления процессом краудсорсинга [5]. Слой данных обеспечивается открытой объектно-реляционной СУБД PostgreSQL. Мобильное приложение для платформы Android доступно в Play Market; исходные коды доступны на GitHub:

- <https://play.google.com/store/apps/details?id=net.russianword.android>;
- <https://github.com/russianwordnet/yarn-android>;
- <https://github.com/mtsar/mtsar>.

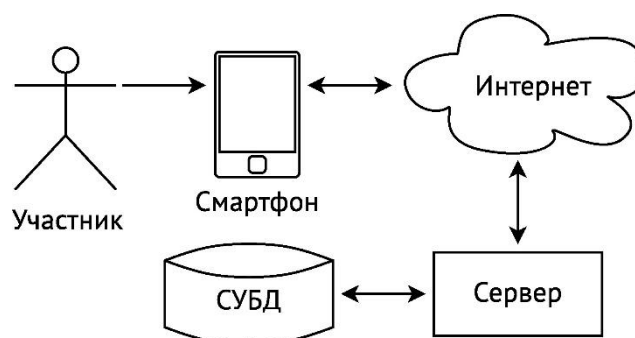


Рис. 1. Архитектурная схема мобильной системы лексических замещений.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-37-00354 мол_а.

1. Gurevych I., Kim J., The People's Web Meets NLP, Springer (2013).
2. Kittur A., Nickerson J.V., Bernstein M. et al., Proc. of CSCW'13, 1301 (2013).
3. Braslavski P., Ustalov D., Mukhin M., Proc. of EACL 2014 (Demo Track), 101 (2014).
4. Biemann C. Language Resources and Evaluation, 47(1), 97 (2013).
5. Усталов Д. Труды Института системного программирования РАН, 27(3), 351 (2015).

ЭФФЕКТИВНОСТЬ ПОКАЗАТЕЛЕЙ WEB-РЕЙТИНГА УНИВЕРСИТЕТОВ 2016

Усманов Р.Ф.^{*}, Неудачин И.Г.

Уральский федеральный университет имени первого Президента России Б.Н. Ельцина,
г. Екатеринбург, Россия

*E-mail: rusman-8@mail.ru

EFFECTIVENESS OF INDICATORS WEBOMETRICS 2016

Usmanov R.F.^{*}, Neudachin I.G.

Ural Federal University, Yekaterinburg, Russia

Performed correlation analysis of World Ranking Web of Universities for January 2016 actual. Estimated impact of four indicators on the final place of the University. The greatest contribution to Overall Rank give an Excellence and Impact indicators.